

Energy Harvesting Communication System with SOC-Dependent Energy Storage Losses

Alessandro Biason and Michele Zorzi
{biasonal,zorzi}@dei.unipd.it

Department of Information Engineering, University of Padova - via Gradenigo 6b, 35131 Padova, Italy

Abstract—The popularity of Energy Harvesting Devices (EHDs) has grown in the past few years, thanks to their capability of prolonging the network lifetime. In reality, EHDs are affected by several inefficiencies, *e.g.*, energy leakage, battery degradation or storage losses. In this work we consider an energy harvesting transmitter with storage inefficiencies. In particular, we assume that when new energy has to be stored in the battery, part of this is wasted and the losses depend upon the current state of charge of the device. This is a practical realistic assumption, *e.g.*, for a capacitor, that changes the structure of the optimal transmission policy. We analyze the throughput maximization problem with a dynamic programming approach and prove that, given the battery status and the channel gain, the optimal transmission policy is deterministic. We derive numerical results for the energy losses in a capacitor and show the presence of a *loop effect* that degrades the system performance if the optimal policy is not considered.

I. INTRODUCTION

One of the most interesting techniques to prolong network lifetime is Energy Harvesting (EH), that allows the nodes of a Wireless Sensor Network (WSN) to refill their batteries with a renewable energy source, *e.g.*, sunlight, vibrations, wind, etc. Differently from traditional networks, where the key aspects are energy conservation and efficiency, in an EH-WSN one of the most important aspects is *energy management*, *i.e.*, in order to improve the communication performance, the Energy Harvesting Devices (EHDs) have to correctly exploit the available resources.

A lot of research has focused on ideal devices [1]–[4]. In reality, EHDs have several inefficiencies, *e.g.*, energy leakage [5], battery degradation [6], imperfect knowledge of the State Of Charge (SOC) [7] or storage losses [8]. In this work we focus on this last problem. In [8], Tutuncuoglu *et al.* assumed that the energy losses can be described with a fixed constant in $[0, 1]$. Our work, instead, assumes that the storage losses *depend upon the current SOC*, which is a realistic assumption, *e.g.*, for a capacitor [9]. Typically, when the battery is depleted or almost fully charged, it is not possible to store a lot of energy (high losses), whereas, when the battery is half-charged, almost all the harvested energy can be stored (low losses). With these considerations, the structure of the traditional optimal transmission policy changes. In this work we find the new optimal policy using a Markov Decision Process approach. Analytically, we prove the threshold structure of the optimal policy, *i.e.*, for every SOC and channel gain, the optimal policy always uses the same transmission power. We present numerical results in

order to describe the properties of the optimal policy. In particular, we notice a *loop effect* (when the battery is almost empty and does not manage to be refilled), that degrades the system performance when the optimal policy is not used. Also, we show that the traditional optimal policy (obtained ignoring storage losses) is strictly sub-optimal in this context. Finally, we describe how the battery size influences the system throughput.

The paper is organized as follows. Section II defines the system model and the optimization problem. In Section III we present the structure of the optimal policy. Section IV presents our numerical results. Finally, Section V concludes the paper.

II. SYSTEM MODEL

We consider an Energy Harvesting Device (EHD) with energy losses in the battery storage process. Time is divided in slots of equal length. In slot k , EHD harvests B_k Joule of energy according to some energy arrival statistic. The harvested energy is divided in two parts: the first fraction, $I_k B_k$, is directly sent into the channel, whereas the second one, $(1 - I_k) B_k$, recharges the device battery. $I_k \in [0, 1]$ is the *power splitting parameter* and is dynamically chosen by the device in each slot k . The battery has a finite size of E_{\max} J and the energy stored in slot k can be exploited only in a later slot. In an ideal situation, all the energy $(1 - I_k) B_k$ is stored in the battery. Here instead, we consider a *real battery* where only a fraction of $(1 - I_k) B_k$ can be stored, while the rest is wasted due to the battery inefficiencies, a behavior that depends upon the current State Of Charge (SOC) E_k of the battery. For example, in a capacitor [9], when SOC is very high or very low, it is not possible to store a lot of energy (very *high losses*), whereas, when SOC is approximately half of the maximum storable energy, then almost all the energy can be stored (very *low losses*). This inefficiency can be described by a function $s(y)$, where y is the current SOC. An example, that was proposed in [9] and will be used as a baseline in this work, is

$$s(y) = 1 - \frac{(y - E_{\max}/2)^2}{\beta_{n.l.}(E_{\max}/2)^2}, \quad (1)$$

which is a symmetric function with respect to $E_{\max}/2$. If the harvested power is ω (assumed constant in a time slot), then a power of $\omega \cdot s(y)$ W can be converted to energy and stored in the battery. When $y = 0$ or $y = E_{\max}$, then only $\omega(1 - 1/\beta_{n.l.})$ W can be converted and $\omega/\beta_{n.l.}$ W are wasted. Ideally, EHD wants to operate at $E_k = E_{\max}/2$ because $s(E_{\max}/2) =$

1. $\beta_{n.l.}$ is a technology parameter strictly greater than one, *e.g.*, $\beta_{n.l.} = 1.05$. When $\beta_{n.l.} \rightarrow \infty$, we have an ideal battery. We remark that the wasted power cannot be used in any other way, *e.g.*, sent directly into the channel, because it is necessary to recharge the battery. In this paper, we use (1), proposed in [9], as an approximate model to describe the energy losses in a capacitor. However, our analytical results do not depend upon the particular structure of $s(y)$, and can be used with other models as well.

The energy drawn from the battery and used for transmission in slot k is $P_k \leq E_k$. Transmitting with power ρ provides a reward $g(\rho, h)$, where h is a parameter related to the channel status. The *reward function* $g(\rho, h)$ is increasing and concave in ρ . An example, that we use as baseline in our numerical evaluation, is $g(\rho, h) = \ln(1 + \Lambda h \rho)$, *i.e.*, $g(\rho, h)$ represents the transmission rate (Λ is an SNR scaling factor). Thus, in slot k the reward is $g(P_k + B_k I_k, H_k)$ (H_k represents the channel gain in slot k , i.i.d. over time and constant in a single slot), *i.e.*, it is given by the sum of the energy drawn from the battery, P_k , and the part of the harvested energy directly sent into the channel, $B_k I_k$. A time slot is structured as follows. At the beginning of the slot, EHD decides P_k and I_k according to its current energy level E_k and to the channel gain H_k . During the slot, the device transmits and harvests energy with constant power. We assume we have only a causal energy arrival information, thus B_k is unknown at the beginning of slot k .

Equation (1) is an instantaneous expression. To convert it to an amount of energy, we have to consider the battery evolution during a slot. Suppose that at the beginning of the slot the energy level is y_0 . After a time t , the energy level is

$$y_t = y_0 + \int_0^t x s(y_\tau) d\tau$$

Note that y_T depends upon 1) the energy stored in the slot, namely x (that is given by the combination of harvested and transmission power), 2) the slot length T and 3) the initial battery level y_0 . We explicitly write y_T as $y_T = y_T(x, y_0)$. In the next we assume a fixed slot length, thus we interchangeably use the notions of energy or power.

The battery evolution is described by the following relation:

$$E_{k+1} = \min\{y_T((1 - I_k)B_k - P_k, E_k), E_{\max}\} \quad (2)$$

Note that also I_k influences the battery evolution. We include P_k in the first argument of $y_T(\cdot, \cdot)$ because, during the slot evolution, the battery level changes also according to the transmission power ($(1 - I_k)B_k - P_k$ represents the stored energy in slot k). The min is used because *battery overflow* may occur.

In this paper we consider energy as measured in discrete energy quanta, and model the system with a discrete Markov Chain (MC). The MC state $(e, h) \in \mathcal{E} \times \mathcal{H} \triangleq \{0, \dots, e_{\max}\} \times \{0, \dots, h_{\max}\}$ corresponds to e energy quanta stored in the battery and to a channel gain h . When e_{\max} energy quanta are stored, the battery is fully charged at E_{\max} . The harvested energy can assume values in $\mathcal{B} = \{0, \dots, b_{\max}\}$. We also apply a *Round* operation to $y_T(\cdot, \cdot)$ and without loss of generality, we require $\text{Round}(y_T(b_{\max}, e')) \geq 1$, where e' is the energy state with the smallest $y_T(b_{\max}, e')$. If this were not

possible, then, when EHD is in state e' , its battery level could not increase. Note that the problem can always be reformulated by changing the notion of energy quanta in order to realize the previous condition. When the number of states is sufficiently large, these assumptions are not critical and the discrete model converges to the continuous one.

A. Optimization

A policy μ , for every MC state, determines the probability of performing a certain action, *i.e.*, μ is a set of random variable pairs $\Theta_{e,h}$ and $\Phi_{e,h}$ (one for every MC state) with joint probability density functions $f(\rho, \iota; e, h)$. Each function maps an action pair $(\rho, \iota) \in \{0, \dots, e\} \times [0, 1]$ (transmission power and power splitting parameter) to a value in \mathbb{R}^+ .

If μ induces a Markov Chain with at most one recurrent class, the long-term average reward does not depend upon the initial state $\mathbf{S}_0 = (E_0, H_0)$. In this case, the long-term average reward G_μ can be expressed as

$$\begin{aligned} G_\mu &= \liminf_{K \rightarrow \infty} \frac{1}{K} \mathbb{E}_{B_k, H_k, P_k, I_k} \left[\sum_{k=0}^{K-1} g(P_k + B_k I_k, H_k) | \mathbf{S}_0 \right] \\ &= \sum_{e \in \mathcal{E}} \pi_\mu(e) j_\mu(e), \end{aligned} \quad (3)$$

where $\pi_\mu(e)$ is the steady state probability of being in state e and $j_\mu(e) \triangleq \mathbb{E}_{B, H} [\mathbb{E}_{\Theta_{e, H}, \Phi_{e, H}} [g(\Theta_{e, H} + B \Phi_{e, H}, H)]]$. $\mathbf{S}_k = (E_k, H_k)$ is the *state of the system* in slot k . B and H are the random variables of the energy arrivals and of the channel gain (i.i.d. over time) with means \bar{b} and \bar{h} . Their pmfs are $p_B(b)$, $b \in \mathcal{B}$ and $p_H(h)$, $h \in \mathcal{H}$, respectively (note that we assume a discrete channel).

In this work we want to find the Optimal Policy (OP) μ^* such that

$$\mu^* = \arg \max_{\mu} G_\mu \quad (4)$$

Therefore, the problem is formulated using the theory of Markov Decision Processes, and can be solved with standard stochastic optimization techniques, *e.g.*, the Policy Iteration Algorithm (PIA) [10]. In Proposition 2 we show that the optimal policy induces a MC with one recurrent class.

III. OPTIMAL POLICY

The aim of this section is to prove a fundamental property of OP.

Theorem 1. *OP is deterministic, i.e., for any given battery level e and channel state h , there exists an optimal action pair (ρ^*, ι^*) that is chosen with probability one to obtain the maximum reward.*

Formally, the previous theorem states that, using OP,

$$f(\rho, \iota; e, h) = \delta_{\rho, \rho^*} \cdot \delta(\iota - \iota^*), \quad \forall e \in \mathcal{E}, \forall h \in \mathcal{H} \quad (5)$$

where $\delta_{\cdot, \cdot}$ is the Kronecker delta function and $\delta(\cdot)$ is the Dirac delta function. In the remainder of this section we prove Theorem 1 in two steps. We first show that in the maximization process of Equation (4) it is sufficient to focus on the second term $j_\mu(e)$ of Equation (3). This is a consequence of

Proposition 1. Later, we apply a Lagrangian approach to show that a deterministic policy maximizes $j_\mu(e)$.

Proposition 1. $\mathbb{P}(E_k = e | \mathbf{S}_0)$ depends upon the policy only through $\mathbb{E}_H[f(\rho, \iota; e, H)]$ for $\rho = 0, \dots, e$, $\forall \iota \in [0, 1]$ and $\forall e \in \mathcal{E}$.

Proof: The proof is by induction on k . At $k = 0$, $\mathbb{P}(E_0 = e | \mathbf{S}_0 = (e_0, h_0))$ is equal to 1 if $e = e_0$ and to 0 otherwise. In this case there is no dependence upon the policy.

Suppose that the thesis is true for k (inductive hypothesis). Using the chain rule, the probability that $E_{k+1} = e'$ given the initial state is

$$\mathbb{P}(E_{k+1} = e' | \mathbf{S}_0) = \sum_{e=0}^{e_{\max}} \mathbb{P}(E_{k+1} = e' | E_k = e) \mathbb{P}(E_k = e | \mathbf{S}_0)$$

Therefore, it is sufficient to show that $\mathbb{P}(E_{k+1} = e' | E_k = e)$ depends upon the policy only through the expectations $\mathbb{E}_H[f(\rho, \iota; e, h)]$. Now, define the set $\mathcal{X}_{e',e} \triangleq \{(\rho, \iota, b) : \min\{y_T((1-\iota)b - \rho, e), e_{\max}\} = e'\}$. We have

$$\begin{aligned} \mathbb{P}(E_{k+1} = e' | E_k = e) &= \sum_{(\rho, \iota, b) \in \mathcal{X}_{e',e}} p_B(b) \sum_{h \in \mathcal{H}} p_H(h) f(\rho, \iota; e, h) \\ &= \sum_{(\rho, \iota, b) \in \mathcal{X}_{e',e}} p_B(b) \mathbb{E}_H[f(\rho, \iota; e, H)] \end{aligned}$$

Thus, $\mathbb{P}(E_{k+1} = e' | E_0)$ depends upon the policy only through the expectations $\mathbb{E}_H[f(\rho, \iota; e, h)]$. ■

For every $e \in \mathcal{E}$, $\rho = 0, \dots, e$ and $\iota \in [0, 1]$, fix a value $\alpha(\rho, \iota; e)$. Consider now the set of policies Ξ such that $\mathbb{E}_H[f(\rho, \iota; e, H)]$ is equal to $\alpha(\rho, \iota; e)$ for every $\mu \in \Xi$. Thanks to Proposition 1, in Equation (3) $\pi_\mu(e)$, $\forall e \in \mathcal{E}$, is the same for every $\mu \in \Xi$. Thus, the maximization of G_μ can be rewritten as $\max_\mu \sum_{e \in \mathcal{E}} \pi_\mu(e) j_\mu(e) = \sum_{e \in \mathcal{E}} \pi_\mu(e) \max_\mu j_\mu(e)$, i.e., G_μ is maximized when $j_\mu(e)$ is maximized for every e . The maximization problem becomes a set of $e_{\max} + 1$ simpler optimization problems.

Focus now on a fixed $e \in \mathcal{E}$. We want to find the optimal $f(\rho, \iota; e, h)$ that maximizes

$$j_\mu(e) = \mathbb{E}_{B,H}[\mathbb{E}_{\Theta_{e,H}, \Phi_{e,H}}[g(\Theta_{e,H} + B\Phi_{e,H}, H)]]$$

subject to the constraints

$$\mathbb{E}_H[f(\rho, \iota; e, H)] = \alpha(\rho, \iota; e), \quad \rho = 0, \dots, e, \quad \iota \in [0, 1]$$

We now want to show that the optimal $f(\rho, \iota; e, h)$ has a deterministic structure using a Lagrangian approach. The Lagrangian function is

$$\begin{aligned} \mathcal{L}(e) &= \mathbb{E}_{B,H} \left[\sum_{\rho=0}^e \left(\int_0^1 f(\rho, \iota; e, H) g(\rho + B\iota, H) d\iota \right) \right] \\ &\quad - \mathbb{E}_H \left[\sum_{\rho=0}^e \int_0^1 v_{\text{th}}(\rho, \iota; e) f(\rho, \iota; e, H) d\iota \right] \\ &= \mathbb{E}_H \left[\sum_{\rho=0}^e \left(\int_0^1 f(\rho, \iota; e, H) \mathbb{E}_B[g(\rho + B\iota, H)] d\iota \right) \right] \\ &\quad - \mathbb{E}_H \left[\sum_{\rho=0}^e \int_0^1 v_{\text{th}}(\rho, \iota; e) f(\rho, \iota; e, H) d\iota \right] \end{aligned}$$

$$= \mathbb{E}_H \left[\sum_{\rho=0}^e \int_0^1 f(\rho, \iota; e, H) (\tilde{g}(\rho, \iota, H) - v_{\text{th}}(\rho, \iota; e)) d\iota \right]$$

where $v_{\text{th}}(\rho, \iota; e)$ is the Lagrange multiplier associated with the constraint $\mathbb{E}_H[f(\rho, \iota; e, H)] = \alpha(\rho, \iota; e)$. Note that we neglected the terms $\alpha(\rho, \iota; e)$ in the definition of $\mathcal{L}(e)$ because they do not contribute in the Lagrangian maximization process. We also introduced a new function $\tilde{g}(\rho, i, h) \triangleq \mathbb{E}_B[g(\rho + B\iota, h)]$ that summarizes the contributions of the energy arrivals. Note that we assume not to know the realization of B but we know its statistics (that is used to compute $\tilde{g}(\rho, \iota, h)$), thus OP depends upon the energy arrival statistics.

The structure of Equation (5) is proved by taking, for a fixed e and h , $(\rho^*, \iota^*) = \arg \max_{(\rho, \iota)_e} (\tilde{g}(\rho, i, h) - v_{\text{th}}(\rho, \iota; e))$.¹

In this section we proved that for every state of the system, there exists an optimal pair that has to be always chosen to obtain the maximum long-term average reward. It is sufficient to find the optimal pair for every system state to define the optimal policy.

We now derive another property of OP, that is useful to perform the numerical evaluation.

Proposition 2. The optimal policy induces a Markov Chain with at most one recurrent class if $p_H(0) > 0$.

Proof: The Markov Chain has two dimensions: the battery and the channel. Since the channel gain is not controlled by the device and changes over time, it is always possible to move along this dimension. We now want to show that it is always possible to go from $e \in \{0, \dots, e_{\max} - 1\}$ to e_{\max} and that the recurrent class is defined by the states that are reachable from e_{\max} .

Since $p_H(0) > 0$, for some k , we have $H_k = 0$ and $E_k < e_{\max}$. In this case, the best policy is to choose $P_k = 0$ and $I_k = 0$, otherwise energy would be wasted without increasing the reward, which is sub-optimal. Thus, there exists a positive probability to go from state $(e, 0)$ to (e', h') , with $e' \geq e + 1$ (no transmissions are performed and new energy quanta arrive) and h' a generic channel state. The reasoning can be iteratively applied to e' until e_{\max} is reached. We showed that it is always possible to increase the state along the battery dimension. The converse (decreasing the state along the battery dimension) in general is not true.

In summary, we have a recurrent class composed of the last part of the MC in the battery dimension and, possibly, a transient class for low e . ■

The proposition guarantees that the problem is well posed, i.e., it is independent of the initial state \mathbf{S}_0 . In reality, the channel gain is a continuous quantity. In this case, we can assume that the device does not transmit for very low channel gains in order to guarantee a minimum SNR level.

IV. NUMERICAL EVALUATION

We use $g(x, H_k) = \ln(1 + \Lambda H_k x)$ as the reward function, thus G_μ represents the throughput of the system. To model

¹The idea behind this result is the following. Consider the maximization over $\xi \in [0, 1]$ of $\xi g_1 + (1 - \xi)g_2$, where g_1 and g_2 are two fixed numbers. The maximum is obtained when $\xi = 1$ if $g_1 > g_2$ and $\xi = 0$ otherwise. In our case ξ is substituted by $f(\rho, \iota; e, H)$ and g_1, g_2 by $\tilde{g}(\rho, i, h) - v_{\text{th}}(\rho, \iota; e)$.

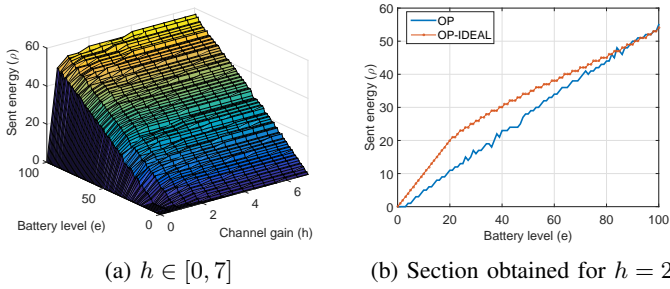


Figure 1: Sent energy $\rho(e, h)$ of OP as a function of the battery state and of the channel gain.

the channel gain, we discretize an exponentially distributed variable (Rayleigh fading). We consider an energy arrivals process described by a truncated geometric r.v. of mean $\bar{b} = 20$ and maximum energy arrivals $b_{\max} = 50$. In this section we want to show the importance of applying the optimal transmission policy to obtain high performance. In particular, we show that the optimal policy computed without considering the inefficiencies of a real battery is strictly sub-optimal.

Several transmission techniques can be applied. Here, we compare three approaches: 1) Optimal Policy (OP) defined in Sec. II-A, 2) optimal policy of an ideal battery (OP-IDEAL) and 3) Greedy Policy (GP). OP-IDEAL is the optimal policy derived considering an ideal battery (without storage losses) and then applied to a real battery (with storage losses). GP is a policy that always poses $I_k = 1$ (the harvested energy is directly sent into the channel and the presence of the battery is neglected). This choice could be interesting for very inefficient or small batteries, but, in general, provides low rewards.

Initially, suppose $I_k = 0$, i.e., the harvested energy cannot be sent directly. In Figure 1a, we show the amount of sent energy ρ of OP for every state of the system when $e_{\max} = 100$, $\Lambda = 0.1$ and $\beta_{n,l} = 1.05$. It can be seen that at $h = 0$ we have $\rho = 0$, as explained in Proposition 2. An interesting point is that the sent energy increases fast when the channel gain is very low and later becomes almost channel independent. This fact can be used to design a low-complexity policy.

In Figure 1b we present a section of the sent energy, obtained when $h = 2$. Here we also represent OP-IDEAL in order to compare the two policies. Note that, when e is not very high, the OP curve is lower than the other. In practice, OP tends to transmit with low powers in order to avoid the low energy states (that have high storage losses). This results (see Figure 2) in $\pi_\mu(e) = 0$ when e is very low (as stated in Proposition 2, there is only one recurrent class). This is particularly important because, due to the energy losses, a *loop effect* may happen. We remember that when e is low or high, a lot of energy is wasted in the recharging process (see the structure of $s(y)$). In particular, focus on the low energy states. If a high power transmission is performed in these states (as per OP-IDEAL), the battery does not manage to be recharged (the transmissions almost compensate the energy arrivals). In practice, the SOC is trapped in the low energy states. Note that also for the high energy states the energy losses are high, but in this case the loop effect is not present because it is always possible to empty the battery, thus moving the energy state to

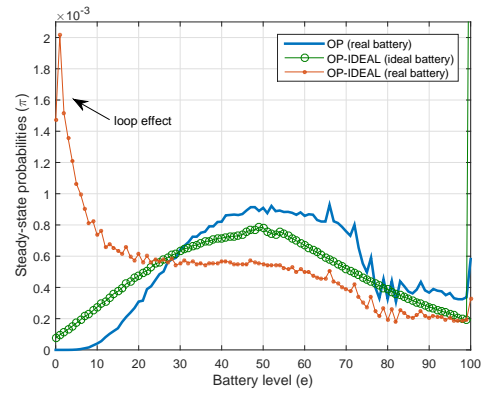


Figure 2: Steady-state probabilities $\pi(e, h)$ of OP and OP-IDEAL (for an ideal as well as a real battery) as a function of the battery state when $h = 2$.

a more advantageous region. These behaviors can be seen in the steady-state probabilities depicted in Figure 2. When we apply OP-IDEAL to the real battery, we can notice the loop effect for the low energy states. This is particularly critical because it degrades the system performance (if e is low, only little energy can be used for transmission, resulting in low rewards) and the device availability (for most of the time EHD is almost in outage). The circled green line represents the steady-state probabilities for an ideal battery. This curve is smoother than the others because the energy arrival distribution is independent of e , whereas in the other cases, for every energy state we have a different energy arrival distribution according to Equation (1). In an ideal battery, the probability of being in the final energy state is high because a lot of energy arrives with respect to the battery size. By increasing the size, the peak decreases. In a real battery instead, since the storage losses are high for high e , the peaks are much smaller.

OP-IDEAL and GP have long-term average throughput equal to 84% and 78% of the throughput of OP. In this case, using OP-IDEAL for a real battery is not convenient, since it has almost the same throughput of GP, that is much simpler to compute and easier to implement.

Figure 3 shows the effects of a finite battery size on the throughput. Focus on the continuous lines. We plot the rewards for OP and OP-IDEAL (for an ideal as well as a real battery). As expected, the OP curve falls between OP-IDEAL in the real and ideal case. When the battery is very small with respect to the energy arrivals, OP-IDEAL and OP are close. This is because a lot of energy arrives and the battery is almost always full. Instead, as e_{\max} increases, the gap between the two policies increases significantly. This is due to the loop effect because the larger the battery, the larger the region where the energy losses are high. Consider the OP curve. It can be seen that it increases slowly for high e_{\max} , e.g., at $e_{\max} = 100$ we have a reward of 0.80 and to increase this value to 0.83 it is necessary to have a much larger battery ($e_{\max} = 140$). This happens because the effects of overflow and outage are reduced with large batteries. Note that the reward saturates and does not grow unbounded.

We now want to show the improvements that can be obtained when $I_k > 0$. Intuitively, when the battery is charged in a high losses region, it is better to directly send the harvested

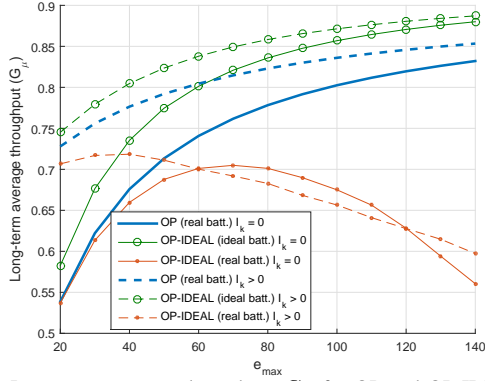


Figure 3: Long-term average throughput G_μ for OP and OP-IDEAL (for an ideal as well as a real battery) as a function of the battery size e_{\max} .

energy into the channel. However, note that, as seen in the previous example, OP already avoids the inefficient regions (the low SOC states are rarely reached).

In Figure 3, with dashed lines, we show the system throughput when I_k is not forced to zero. For OP and OP-IDEAL for an ideal battery, as expected, the curves are above their counterparts with $I_k = 0$. Instead, for OP-IDEAL for a real battery, it may be better not to use $I_k > 0$ in some regions. Thus, as can be seen, even with $I_k > 0$, OP-IDEAL is strictly sub-optimal when applied to a real battery. Intuitively, allowing to directly send the harvested energy does not provide a great improvement when the battery is large: in this case it is sufficient to operate in a region with low energy losses in order to efficiently store the harvested energy and use it later. This is not possible in a small battery, and in this case directly sending the harvested energy may provide a significant reward improvement (left part of Figure 3).

With the parameters of Figure 1a, when we find OP with $I_k > 0$, the reward becomes 0.83 (instead of 0.80). Even if the relative improvement is small, note that, with $I_k = 0$, it is necessary to have a much larger battery to obtain a reward of 0.83 ($e_{\max} = 140$). Thus, allowing $I_k > 0$ permits to obtain higher rewards without increasing the battery size. In Figure 4 we show the power splitting parameter ι for every battery state and channel gain. When SOC is high or low, the energy losses are high, thus in these cases it is better to directly send the harvested energy. As e moves toward $e_{\max}/2$, also $\iota(e, h)$ decreases until the mid region where $\iota(e, h) \approx 0.5$. In this case it is convenient to store the energy for future use because the losses are very low and, at the same time, to directly send energy into the channel to obtain an immediate reward. When the SOC is low, the battery can still be charged by exploiting the states where the channel gain is zero.

V. CONCLUSIONS

We studied an Energy Harvesting Device with imperfect energy storage capabilities. During the energy saving process, part of the incoming energy is wasted and this effect depends upon the current state of charge of the device. We proved that, for given battery status and channel gain, the Optimal Policy is deterministic. We also showed that the underlying MC induced by OP has at most one recurrent class. These results were used to develop the numerical evaluation, where we studied how OP

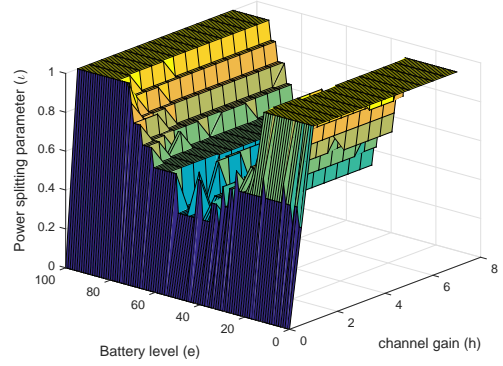


Figure 4: Power splitting parameter $\iota(e, h)$ for OP as a function of the battery state e and of the channel state h .

influences the performance of an EHD with a capacitor. We found that the energy losses degrade the system performance, making the device fall into a region where the loop effect may be present. Finding the optimal policy in this context is useful in order to avoid the loop effect and to restore high rewards. We noticed that the reward increases slowly with the battery size and is upper bounded. Finally, we showed that allowing $I_k > 0$ is equivalent, in terms of reward, to using a much larger battery, thus implementing the power splitting mechanism is convenient when the battery is small.

Possible extensions consist in the relaxation of some hypothesis, *e.g.*, those about perfect SOC and channel knowledge. Also, the circuitry costs can be taken into account. Moreover, it would be interesting to study the effects of imperfect energy storage with other functions (in addition to (1)).

REFERENCES

- [1] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Trans. Wireless Comm.*, vol. 8, no. 2, pp. 547–551, Feb. 2009.
- [2] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Comm.*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.
- [3] O. Ozel and S. Ulukus, "Information-theoretic analysis of an energy harvesting communication system," in *21st International Symposium on Personal, Indoor and Mobile Radio Communications Workshops (PIMRC Workshops)*, Sep. 2010, pp. 330–335.
- [4] —, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6471–6483, Jun. 2012.
- [5] B. Devillers and D. Gunduz, "A general framework for the optimization of energy harvesting communication systems with battery imperfections," *IEEE J. of Commun. and Netw.*, vol. 14, no. 2, pp. 130–139, Apr. 2012.
- [6] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Energy management policies for harvesting-based wireless sensor devices with battery degradation," *IEEE Trans. Commun.*, vol. 61, no. 12, Nov. 2013.
- [7] N. Michelusi, L. Badia, and M. Zorzi, "Optimal transmission policies for energy harvesting devices with limited state-of-charge knowledge," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 3969–3982, Sep. 2014.
- [8] K. Tutuncuoglu, A. Yener, and S. Ulukus, "Optimum policies for an energy harvesting transmitter under energy storage losses," *IEEE J. on Selected Areas in Commun.*, vol. 33, no. 3, pp. 467–481, Jan. 2015.
- [9] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," *IEEE Trans. Mobile Computing*, vol. 12, no. 9, pp. 1853–1865, Jul. 2013.
- [10] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, Belmont, Massachusetts, 2005.